



J.W. Price, 949/261.8433

Hideki Fujimori et al

S.N. 09/821,525

日 本 国 特 許 庁

PATENT OFFICE DAK1-13030
JAPANESE GOVERNMENT

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日

Date of Application:

2000年 3月29日

出 願 番 号

Application Number:

特願2000-090337

出 願 人

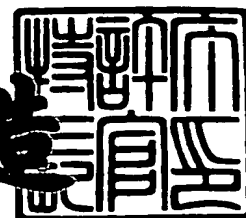
Applicant (s):

松下電器産業株式会社

2001年 1月19日

特許庁長官
Commissioner,
Patent Office

及 川 耕 造



出証番号 出証特2000-3112873

【書類名】 特許願

【整理番号】 2022520064

【提出日】 平成12年 3月29日

【あて先】 特許庁長官殿

【国際特許分類】 G06F 13/00

【発明者】

【住所又は居所】 広島県東広島市鏡山3丁目10番18号 株式会社松下
電器情報システム広島研究所内

【氏名】 藤森 秀樹

【発明者】

【住所又は居所】 広島県東広島市鏡山3丁目10番18号 株式会社松下
電器情報システム広島研究所内

【氏名】 澄田 哲二

【特許出願人】

【識別番号】 000005821

【氏名又は名称】 松下電器産業株式会社

【代理人】

【識別番号】 100097445

【弁理士】

【氏名又は名称】 岩橋 文雄

【選任した代理人】

【識別番号】 100103355

【弁理士】

【氏名又は名称】 坂口 智康

【選任した代理人】

【識別番号】 100109667

【弁理士】

【氏名又は名称】 内藤 浩樹

【手数料の表示】

【予納台帳番号】 011305

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9809938

【書類名】 明細書

【発明の名称】 データ中継システム

【特許請求の範囲】

【請求項 1】 複数クライアント端末からのインターネット上のコンテンツ取得要求を Web サーバに中継し、コンテンツ取得要求のレスポンスとして Web サーバから受信したコンテンツをキャッシュに蓄積し、コンテンツをクライアント端末に転送するデータ中継システムであって、

クライアント端末－データ中継システム間の通信状況、および、データ中継システム－Web サーバ間の通信状況により、Web サーバからのコンテンツ受信の優先順位を変更するデータ受信制御手段を備える、データ中継システム。

【請求項 2】 複数クライアント端末からのインターネット上のコンテンツ取得要求を Web サーバに中継し、コンテンツ取得要求のレスポンスとして Web サーバから受信したコンテンツをキャッシュに蓄積し、コンテンツをクライアント端末に転送するデータ中継システムであって、

クライアント端末－データ中継システム間の通信状況、および、データ中継システム－Web サーバ間の通信状況により、クライアント端末に対して、Web サーバから受信したコンテンツの転送開始タイミングを変更するデータ送信制御手段を備える、データ中継システム。

【請求項 3】 前記データ中継システムは、クライアント端末の能力を取得する端末能力取得手段をさらに備え、前記データ受信制御手段は、クライアント端末の能力を元にコンテンツ受信の優先順位を決定することを特徴とする請求項 1 に記載のデータ中継システム。

【請求項 4】 前記データ中継システムは、クライアント端末へのコンテンツ転送状況を監視する転送状況監視手段をさらに備え、前記データ受信制御手段は、クライアント端末へのコンテンツ転送状況を元にコンテンツ受信の優先順位を決定することを特徴とする請求項 1 に記載のデータ中継システム。

【請求項 5】 前記データ中継システムは、Web サーバへのリクエスト転送からコンテンツ取得までのレスポンスを監視するレスポンス監視手段をさらに備え、前記データ受信制御手段は、Web サーバのレスポンスを元にコンテンツ受

信の優先順位を決定することを特徴とする請求項 1 に記載のデータ中継システム

【請求項 6】 前記データ中継システムは、クライアント端末の能力を取得する端末能力取得手段をさらに備え、前記データ送信制御手段は、クライアント端末の能力を元に W e b サーバから受信したコンテンツの転送開始タイミングを変更することを特徴とする請求項 2 に記載のデータ中継システム。

【請求項 7】 前記データ中継システムは、W e b サーバへのリクエスト転送からコンテンツ取得までのレスポンスを監視するレスポンス監視手段をさらに備え、前記データ送信制御手段は、W e b サーバのレスポンスを元に W e b サーバから受信したコンテンツの転送開始タイミングを変更することを特徴とする請求項 2 に記載のデータ中継システム。

【発明の詳細な説明】

【 0 0 0 1 】

【発明の属する技術分野】

本発明は、インターネットなどの情報通信ネットワークにおいて、オーディオ、ビデオなどのストリーム・データを W e b サーバから取得し、クライアント端末に転送するデータ転送システムに関する。

【 0 0 0 2 】

【従来の技術】

従来、インターネット上の W e b サーバからコンテンツを取得し、クライアント端末に転送する方法として、クライアント端末－W e b サーバ間の適切な箇所にプロキシ・サーバを置く技術が一般的である。

【 0 0 0 3 】

一般的なプロキシ・サーバの構成を図 1 4 を参照して説明する。図 1 4 を参照して、クライアント端末 1 4 0 1 からのリクエスト受付を行うリクエスト受付部 1 4 0 4 と、W e b サーバにリクエストを転送するリクエスト転送部 1 4 0 5 と、W e b サーバからのコンテンツを受信するデータ受信部 1 4 0 6 と、W e b サーバからのコンテンツ受信を W e b サーバとのセッション管理している順番で決定するデータ受信制御部 1 4 0 7 と、コンテンツを蓄積するキャッシュ保存部 1 4 0 9 と、クライアント端末にコンテンツを送信するデータ送出部 1 4 0 8 とを

有している。

【0004】

また、前記一般的なプロキシ・サーバの動作の概略は次のようになる。リクエスト受付部1404は、クライアント端末1401からデータの識別子であるURL (Uniform Resource Locator) によるデータ取得のリクエストを受信し、リクエスト転送部1405に転送する。リクエスト転送部1405は、リクエストをWebサーバに送信する。データ受信制御部1407は、Webサーバとのセッションの管理順に従って、データ受信の順番を決定する。データ受信部1406は、データ受信制御部1407が決定した順番に従って、Webサーバからコンテンツを受信し、キャッシュ保存部1409に保存する。データ送出部1408は、キャッシュに保存されたコンテンツをクライアント端末1401に送信する。

【0005】

以上のような、一般的なプロキシ・サーバにおいて、クライアント端末からリクエストされたコンテンツが音声、動画などのストリーム・データの場合、クライアント端末-プロキシ・サーバ間の転送レートが遅ければ、クライアント端末に対するストリーム・データ送信で遅延が生じ、リアルタイム性が保証できなくなる。

【0006】

この課題を解決するための方法として、クライアント端末で、ストリーム・データの再生を開始するまでに、リクエストしたコンテンツをWebサーバからある程度先読みする方法が考えられるが、この方法の場合、各クライアント端末毎にコンテンツを先読みするために多くのメモリを必要とするため、あまり効率的であるとは言えない。そこで、プロキシ・サーバでこの課題を解決する方法を見てもみると、特開2000-13779号公報の方法が該当すると考えられる。特開2000-13779号公報では、クライアント端末-プロキシ・サーバ間の転送レートにより、Webサーバから取得したストリーム・データの品質を落とすことで、ストリーム・データのサイズを小さくし、クライアント端末へのストリーム・データ送信のリアルタイム性を保証している。

【 0 0 0 7 】

【発明が解決しようとする課題】

しかしながら、上記方法では、プロキシ・サーバーWebサーバ間の転送レートが考慮されておらず、Webサーバからのストリーム・データ受信が一定時間滞った場合、プロキシ・サーバからクライアント端末に対するストリーム・データ送信において、データ途切れが発生してしまう。

【 0 0 0 8 】

図2を参照して、データ途切れの発生する状況を説明する。縦軸はストリーム・データサイズで、横軸は時間の経過を示す。Webサーバからのデータ受信が時間T1で滞るとすると、Webサーバから受信するストリーム・データサイズは201で表すことができる。また、プロキシ・サーバはWebサーバからのストリーム・データの受信と平行して、クライアント端末へのストリーム・データの送信を行っているため、クライアント端末への送信済みストリーム・データサイズは202で表すことができる。201と202で囲まれた部分の203がクライアント端末に未送信のストリーム・データとなり、時間T2でクライアント端末へ未送信のストリーム・データがアンダーフローを起こし、データ途切れが発生する。

【 0 0 0 9 】

クライアント端末へ未送信のストリーム・データのアンダーフローは、クライアント端末ープロキシ・サーバ間の転送レートが速い場合と、プロキシ・サーバーWebサーバ間の転送レートが遅い場合に、同様に発生してしまう。

【 0 0 1 0 】

ここで、複数のクライアント端末とのセッションを有するプロキシ・サーバにおいて、各セッション毎の状態は、上記アンダーフローを起こすセッションと、未送信のストリーム・データ203に余裕のあるセッションが共存していることが考えられる。

【 0 0 1 1 】

このため、本発明の目的は、未送信のストリーム・データが一定時間アンダーフローを起こさないようなタイミングでクライアント端末に対するストリーム・

データ送信を開始し、未送信のストリーム・データに余裕のないセッションを優先して、We bサーバからのストリーム・データの受信処理を行うことで、データ途切れが起きにくいデータ中継システムを提供することである。

【 0 0 1 2 】

【課題を解決するための手段】

第 1 の発明は、複数クライアント端末からのインターネット上のコンテンツ取得要求をWe bサーバに中継し、コンテンツ取得要求のレスポンスとしてWe bサーバから受信したコンテンツをキャッシュに蓄積し、コンテンツをクライアント端末に転送するデータ中継システムであって、

クライアント端末ーデータ中継システム間の通信状況、および、データ中継システムーWe bサーバ間の通信状況により、We bサーバからのコンテンツ受信の優先順位を変更するデータ受信制御手段を備える。

【 0 0 1 3 】

上記のように、本発明によれば、クライアント端末ーデータ中継システム間の通信状況、および、データ中継システムーWe bサーバ間の通信状況により、We bサーバからのコンテンツ受信の優先順位を変更するため、通信状況が変動した場合に、クライアント端末に送信するコンテンツのデータ途切れが起こり難くなる。

【 0 0 1 4 】

また、本発明は、複数クライアント端末からのインターネット上のコンテンツ取得要求をWe bサーバに中継し、コンテンツ取得要求のレスポンスとしてWe bサーバから受信したコンテンツをキャッシュに蓄積し、コンテンツをクライアント端末に転送するデータ中継システムであって、

クライアント端末ーデータ中継システム間の通信状況、および、データ中継システムーWe bサーバ間の通信状況により、クライアント端末に対して、We bサーバから受信したコンテンツの転送開始タイミングを変更するデータ送信制御手段を備える。

【 0 0 1 5 】

上記のように、本発明によれば、クライアント端末ーデータ中継システム間の

通信状況、および、データ中継システム－Webサーバ間の通信状況により、Webサーバから受信したコンテンツの転送開始タイミングを変更するため、クライアント端末に対し、コンテンツを送信を開始する時点で、一定時間分のコンテンツを保証することが可能となり、通信状況が変動した場合に、クライアント端末に送信するコンテンツのデータ途切れが起こり難くなる。

【 0 0 1 6 】

また本発明は、クライアント端末の能力を取得する端末能力取得手段をさらに備え、データ受信制御手段は、クライアント端末の能力を元にコンテンツ受信の優先順位を決定することを特徴とする。

【 0 0 1 7 】

上記のように、本発明によれば、第1の発明におけるデータ受信制御手段が、クライアント端末の能力を元にコンテンツ受信の優先順位を決定するため、クライアント端末の最大能力でコンテンツ送信した場合にも、データ途切れが起こり難くなる。

【 0 0 1 8 】

また本発明は、クライアント端末へのコンテンツ転送状況を監視する転送状況監視手段をさらに備え、データ受信制御手段は、クライアント端末へのコンテンツ転送状況を元にコンテンツ受信の優先順位を決定することを特徴とする。

【 0 0 1 9 】

上記のように、本発明によれば、データ受信制御手段が、クライアント端末へのコンテンツ転送状況を元にコンテンツ受信の優先順位を決定するため、クライアント端末の通信状況の変動に対応しながらも、データ途切れが起こり難くなる。

【 0 0 2 0 】

また本発明は、Webサーバへのリクエスト転送からコンテンツ取得までのレスポンスを監視するレスポンス監視手段をさらに備え、データ送信制御手段は、Webサーバのレスポンスを元にWebサーバから受信したコンテンツの転送開始タイミングを変更することを特徴とする。

【 0 0 2 1 】

上記のように、本発明によれば、データ送信制御手段が、Webサーバのレスポンスを元にWebサーバから受信したコンテンツの転送開始タイミングを変更するため、Webサーバからのコンテンツ受信にかかる遅延時間を含んだ一定時間分のコンテンツを保証することが可能となり、通信状況が変動した場合に、クライアント端末に送信するコンテンツのデータ途切れが起こり難くなる。

【0022】

また本発明は、クライアント端末の能力を取得する端末能力取得手段をさらに備え、データ送信制御手段は、クライアント端末の能力を元にWebサーバから受信したコンテンツの転送開始タイミングを変更することを特徴とする。

【0023】

上記のように、本発明によれば、データ送信制御手段が、クライアント端末の能力を元にWebサーバから受信したコンテンツの転送開始タイミングを変更するため、能力の高いクライアント端末に対しても、一定時間分のコンテンツを保証することが可能となり、通信状況が変動した場合に、クライアント端末に送信するコンテンツのデータ途切れが起こり難くなる。

【0024】

【発明の実施の形態】

以下、本発明の実施形態について図面を参照しながら説明する。

【0025】

図1は、本発明の一実施形態に係るデータ中継システムの構成を示すブロック図である。図1において、本発明の一実施形態に係るプロキシ・サーバは公衆網2および、インターネット12と接続している。

【0026】

まず、本発明の一実施形態に係るプロキシ・サーバの詳細な構成を説明する。

【0027】

図1を参照して、インターネット・プロキシ・サーバはリクエスト受付部4と、リクエスト転送部5と、データ受信部6と、データ送受信制御部7と、データ送出部8と、転送状況監視部9と、レスポンス監視部10と、キャッシュ11とを備える。

【 0 0 2 8 】

リクエスト受付部 4 は、クライアント端末 1 からのリクエストおよび、端末能力を受け付け、リクエストされたコンテンツがキャッシュ 1 1 に存在する場合、データ送出部 8 に対し、クライアント端末 1 への該当コンテンツの送信を要求する。リクエスト受付部 4 は、リクエストされたコンテンツがキャッシュ 1 1 に存在しない場合、リクエストをリクエスト転送部 5 に転送する。リクエスト転送部 5 は、リクエスト受付部 4 から転送されたリクエストを W e b サーバ 1 3 に転送する。

【 0 0 2 9 】

データ送受信制御部 7 は、図 2 のキャッシュ管理テーブルを管理し、データ受信部 6 の U R L 毎の受信処理の優先順位付けを行い、また、データ送出部 8 のデータ送出開始のタイミングを判定する。データ受信部 6 は、U R L 毎の優先順位に従って W e b サーバ 1 3 から受信して、受信したコンテンツをキャッシュ 1 1 に保存する。データ送出部 8 は、クライアント端末 1 のリクエストに対応するキャッシュ 1 1 内のコンテンツをクライアント端末 1 に送出する。

【 0 0 3 0 】

転送状況監視部 9 は、データ送出部 8 のデータ送出状況を監視し、データ送受信制御部 7 に通知する。レスポンス監視部 1 0 は、データ受信部 6 のデータ受信状況を監視する。キャッシュ 1 1 は、データ受信部 6 が W e b サーバ 1 3 から受信したコンテンツを保存する記憶装置である。

【 0 0 3 1 】

なお、転送状況監視部 9 のデータ送出状況を監視する形態としては、例えば、データを送信するのに要した時間と、送信したデータサイズから検出する形態が考えられる。

【 0 0 3 2 】

また、レスポンス監視部 1 0 のデータ受信状況を監視する形態としては、リクエスト転送部 5 の、リクエスト転送以降に、データ受信部 6 がデータを受信するまでに要した時間により検出する形態が考えられる。

【 0 0 3 3 】

以下、本発明の一実施形態に係るプロキシ・サーバが行うクライアント端末1からのリクエストに対応するコンテンツをWebサーバ13から取得し、クライアント端末1aに送出する際のキャッシュ制御方法について、図1、図3～図5を参照して具体的に説明を行う。

【0034】

最初に、図3を参照して、各データの関係の説明する。図3は、データ送受信制御部7がWebサーバ13からの受信処理の優先順位を判定するために、Webサーバ13から受信したコンテンツで、クライアント端末1に対して、一定時間分の送信データサイズ（以下、未送信バッファサイズという）を制御する際に用いる管理データテーブル（以下、キャッシュ管理テーブル300という）である。また、このキャッシュ管理テーブル300は、データ送受信制御部7が、データ送出部8に対して、クライアント端末1へのコンテンツ送信開始のタイミングを指示する際にも使用する。

【0035】

端末ID301は、クライアント端末を特定する一意の識別子である。URL302は、クライアント端末からリクエストされたコンテンツのURLである。WebサーバセッションID303は、前記URL302のリクエストを転送するWebサーバ13とのセッションを特定する一意の識別子である。

【0036】

最大端末転送レート304は、リクエスト受付部4がクライアント端末1とセッション確立時に端末情報として受取る端末の通信速度の最大値である。端末転送レート305は、転送状況監視部9が取得する単位時間当りのクライアント端末1への転送レートの実測値である。Webサーバレスポンス306は、リクエスト転送部5が、Webサーバ13にリクエストを送信して、データ受信部6が、結果を受信するまでの遅延時間である。未送信バッファサイズ307は、クライアント端末1に対して、コンテンツを一定時間送出するために必要なデータサイズである。

【0037】

未送信データサイズ308は、Webサーバ13から受信済みのコンテンツで

クライアント端末1に対して、送出していないコンテンツのサイズである。送信済みデータサイズ309は、Webサーバ13から受信したコンテンツで、データ送出部8が、クライアント端末1に送出済みのデータサイズである。コンテンツサイズ310は、クライアント端末1からリクエストされたURLのコンテンツのデータサイズである。

【0038】

ここで、未送信バッファサイズ307は、最大端末転送レート304、端末転送レート305、Webサーバレスポンス306により、算出されるものである。

【0039】

次に、図4を参照して、各データの関係性を説明する。図4は、レスポンス監視部10が、Webサーバ13のレスポンス時間を判定するために用いる管理テーブル（以下、レスポンス監視テーブル400という）である。

【0040】

WebサーバセッションID401は、Webサーバ13とのセッションを識別するための一意の識別子である。リクエスト送信時刻402は、リクエスト転送部5が、Webサーバに対し、リクエストを送信した時刻である。応答受信時刻403は、データ受信部6が、Webサーバからデータ受信した時刻である。コンテンツサイズ404は、クライアント端末1からリクエストされたURLのコンテンツのデータサイズである。

【0041】

上述したデータ関係において、データ送受信制御部7がキャッシュ管理テーブル300の各フィールドを更新し、更新した未送信バッファサイズ307を元に、データ受信部6のコンテンツ受信制御を行う方法、および、データ送出部8のデータ送出制御を行う方法を説明する。

【0042】

図5を参照して、リクエスト受付部4は、クライアント端末1からのコンテンツ取得のリクエストを受信する（S501）。S501のリクエスト受け付け時に、リクエスト受付部4は、リクエストと共に、端末IDおよび、最大端末転送レ

ートも同時に受信する。なお、この端末ID、および、最大端末転送レートは、プロキシ・サーバに予め登録しておき、クライアント端末1からのリクエスト受信時にリクエスト受付部4が検索した結果を使用してもよい。

【0043】

リクエスト受付部4は、リクエストされたURLに対応したコンテンツがキャッシュに存在するか検索し、キャッシュに存在している場合、通常のプロキシ・サーバの動作と同様に、データ送出部8にデータ送出開始を要求し、データ送出部8は、キャッシュ11のコンテンツをクライアント端末1に送出する。

【0044】

リクエストされたURLに対応したコンテンツがキャッシュに存在しない場合、リクエスト受付部4は、リクエストと、端末IDをリクエスト転送部5に転送する(S502)。また、リクエスト受付部4は、クライアント端末1の端末IDと、最大端末転送レート、リクエストURLをデータ送受信制御部7に通知する(S503)。

【0045】

データ送受信制御部7は、キャッシュ管理テーブル300に新規に項目を確保し、S503でリクエスト受付部4から通知された情報をそれぞれ、端末ID301、URL302、最大端末転送レート304に保持し、最大端末転送レート304により、未送信バッファサイズの初期値を算出し、未送信バッファサイズ307に保持する。この未送信バッファサイズの初期値の算出方法は、クライアント端末1に対し、最大端末転送レート304で、一定時間コンテンツを送出し続けるのに必要なデータサイズである。

【0046】

例えば、最大端末転送レート9600bps(bit per second)のクライアント端末1に対し、10秒間のストリームデータを送出し続けるために必要なデータサイズは、「 $(9600/8) \times 10 = 12000 \text{ byte}$ 」となる。

【0047】

S502により、リクエスト転送部5は、リクエストされたURLのコンテン

ツサイズの取得要求をWebサーバ13に送信し(S504)、Webサーバ13とのセッションを識別するためのセッションID、コンテンツサイズ取得要求を送信した時刻をレスポンス監視部10に通知する(S505)。レスポンス監視部は、S505でリクエスト転送部5から通知された情報を、レスポンス監視テーブル400のWebサーバセッションID401、リクエスト送信時刻402に保持する。

【0048】

データ受信部6は、S504に対する応答を、Webサーバ13から受信し(S506)、Webサーバ13とのセッションID、コンテンツサイズ、応答受信時刻をレスポンス監視部10に通知する(S507)。

【0049】

レスポンス監視部10は、データ受信部6からS507で通知された情報を、レスポンス監視テーブル400のWebサーバセッションID401が一致する項目の応答受信時刻403、コンテンツサイズ404に保持し、Webサーバのレスポンス時間を算出する。このWebサーバのレスポンス時間の算出方法は、応答受信時刻403とリクエスト送信時刻402の差分の時間である。

【0050】

例えば、応答受信時刻403が、 17000 msec で、リクエスト送信時刻402が、 15000 msec の場合、Webサーバのレスポンス時間は、「 $17000 - 15000 = 2000\text{ msec}$ 」となる。また、レスポンス監視部10は、WebサーバセッションID401、コンテンツサイズ404、および、算出したWebサーバのレスポンス時間をデータ送受信制御部7に通知する(S508)。

【0051】

データ送受信制御部7は、レスポンス監視部10からS508で通知された情報を、キャッシュ管理テーブル300のWebサーバセッションID303の一致する項目のWebサーバレスポンス306、コンテンツサイズ310に保持し、未送信バッファサイズ307を再計算する。

【0052】

この未送信バッファサイズ307の再計算方法は、クライアント端末1に対し、最大端末転送レート304で、一定時間コンテンツを送出し続けるのに必要なデータサイズに、クライアント端末1に対し、最大端末転送レート304で、Webサーバ13のレスポンス時間分コンテンツを送出し続けるのに必要なデータサイズを加えたものである。

【0053】

例えば、Webサーバのレスポンス時間が2秒で、最大端末転送レートが9600bpsのクライアント端末1に対し、10秒間のストリームデータを送出し続けるために必要なデータサイズは、「 $(9600/8) \times (10+2) = 14400 \text{ byte}$ 」となる。ただし、未送信バッファサイズ307がコンテンツサイズ310以上になる場合、未送信バッファサイズ307をコンテンツサイズ310と同じ値にする。

【0054】

データ送受信制御部7は、リクエスト転送部5に対し、コンテンツ取得のリクエストを条件付きでWebサーバセッションID303のWebサーバ13に送信するよう要求する(S509)。この条件付リクエストは、条件として、コンテンツの先頭から、未送信バッファサイズ307に保持しているサイズ分のコンテンツ取得を指定する。

【0055】

リクエスト転送部5は、データ送受信制御部7から指定された条件付リクエストをWebサーバ13に送信し(S510)、WebサーバセッションID、リクエスト送信時刻をレスポンス監視部10に通知する(S511)。

【0056】

レスポンス監視部は、S505でリクエスト転送部5から通知された情報を、レスポンス監視テーブル400のWebサーバセッションID401の一致する項目の、リクエスト送信時刻402に保持し、応答受信時刻403をクリアする。

【0057】

データ受信部6は、S510に対する応答としてコンテンツの一部を、Web

サーバ13から受信し（S512）、キャッシュ11に保持する。また、データ受信部6は、WebサーバセッションID、応答受信時刻、今回受信したコンテンツのデータサイズをレスポンス監視部10に通知する（S513）。

【0058】

レスポンス監視部10は、データ受信部6からS513で通知された情報を、レスポンス監視テーブル400のWebサーバセッションID401が一致する項目の応答受信時刻403に保持し、Webサーバのレスポンス時間を算出する。このWebサーバのレスポンス時間の算出方法は、応答受信時刻403とリクエスト送信時刻402の差分の時間である。また、レスポンス監視部10は、WebサーバセッションIDと、Webサーバレスポンスと、S512でデータ受信部6から通知された受信したコンテンツのデータサイズをデータ送受信制御部7に通知する（S514）。

【0059】

データ送受信制御部7は、S514で通知された情報で、キャッシュ管理テーブル300のセッションID303の一致する項目のWebサーバレスポンス306、未送信データサイズ308、送信済みデータサイズ309を更新し、未送信バッファサイズ307を再計算する。この未送信バッファサイズ307の再計算方法は、クライアント端末1に対し、最大端末転送レート304で、一定時間コンテンツを送出し続けるのに必要なデータサイズに、クライアント端末1に対し、最大端末転送レート304で、Webサーバ13のレスポンス時間分コンテンツを送出し続けるのに必要なデータサイズを加えたものである。

【0060】

データ送受信制御部7は、未送信データサイズ308と、送信済みデータサイズ309の合計サイズをコンテンツサイズ310と比較し、Webサーバ13aからのコンテンツ受信が完了しているか判定する。この判定において、Webサーバ13aからのコンテンツ受信が完了していない場合、データ送受信制御部7は、クライアント端末1に対して、一定時間分の未送信データが確保されているかを判定するために、未送信データサイズ308と未送信バッファサイズ307を比較し、未送信データサイズ308が未送信バッファサイズ307以上確保さ

れていない場合、S 5 0 9 から繰り返し処理する。

【 0 0 6 1 】

データ送受信制御部 7 は、未送信データサイズ 3 0 8 が未送信バッファサイズ 3 0 7 以上確保されている場合、データ送出部 8 に端末 I D 3 0 1 のクライアント端末 1 に対して、U R L 3 0 2 のコンテンツの送出開始を要求する（S 5 1 5 ）。

【 0 0 6 2 】

S 5 1 5 により、データ送出部 8 は、クライアント端末 1 a に対し、キャッシュ 1 1 のコンテンツの送出を開始し（S 5 1 6 ）、転送状況監視部 9 にデータ送出に要した時間と、送信したコンテンツのデータサイズを通知する（S 5 1 7 ）。

【 0 0 6 3 】

転送状況監視部 9 は、S 5 1 7 により通知された情報を元にクライアント端末 1 a の端末転送レートを算出する。この端末転送レートの算出方法は、単位時間当りのクライアント端末 1 a への送出データサイズである。転送状況監視部 9 は、端末 I D と、算出した端末転送レート、送信したコンテンツのデータサイズをデータ送受信制御部 7 に通知する（S 5 1 8 ）。

【 0 0 6 4 】

データ送受信制御部 7 は、S 5 1 8 で通知された情報で、キャッシュ管理テーブル 3 0 0 の端末 I D 3 0 1 の一致する項目の端末転送レート 3 0 5、未送信データサイズ 3 0 8、送信済みデータサイズ 3 0 9 で更新し、未送信バッファサイズ 3 0 7 を再計算する。この未送信バッファサイズ 3 0 7 の再計算方法は、クライアント端末 1 に対し、端末転送レート 3 0 5 で、一定時間コンテンツを送出し続けるのに必要なデータサイズに、クライアント端末 1 に対し、端末転送レート 3 0 5 で、W e b サーバ 1 3 のレスポンス時間分コンテンツを送出し続けるのに必要なデータサイズを加えたものである。

【 0 0 6 5 】

また、データ送受信制御部 7 は、キャッシュ 1 1 に、クライアント端末 1 に対して未送信のコンテンツが一定時間分確保できているかを判定するために、未送

信データサイズ308と未送信バッファサイズ307を比較し、未送信データサイズ308が未送信バッファサイズ307以上確保されていない場合、リクエスト転送部5に対し、コンテンツ取得のリクエストを条件付きでWebサーバセッションID303のWebサーバ13に送信するよう要求する(S519)。この条件付リクエストは、条件として、取得済みコンテンツの続きから、未送信バッファサイズ307に保持しているサイズ分のコンテンツ取得を指定する。

【0066】

ここで、データ送受信制御部7は、Webサーバ13からのコンテンツ受信を行っているデータ受信部6において、コンテンツの受信処理に余裕がある場合、例えば、キャッシュ管理テーブル300の全ての未送信データサイズ308が十分確保されている場合では、キャッシュ管理テーブル300に保持している項目毎にコンテンツ取得のリクエストをWebサーバ13に送信するよう、リクエスト転送部に要求するなどの方法で、Webサーバからのコンテンツの受信処理を制御する。以降の処理は、S509からの繰り返しとなる。

【0067】

次に、本発明の実施形態を具体的な数値を用い、図1、図3～図4、図6～図13を参照しながら説明する。

【0068】

図6～図9は、クライアント端末1aからリクエストを受付けた際に、図4のレスポンス監視テーブル400の一項目が処理経過により、更新される内容を表記したものである。

【0069】

図10～図13は、クライアント端末1aからリクエストを受付けた際に、図3のキャッシュ管理テーブル300の一項目が処理経過により、更新される内容を表記したものである。

【0070】

図1を参照して、クライアント端末1aからコンテンツ取得のリクエストを受付けた場合の動作を説明する。

【0071】

この時、図1のキャッシュ11、および、図3のキャッシュ管理テーブル300の各項目301~310、および、図4のレスポンス監視テーブル400の各項目401から404に登録済みのデータはないものとする。

【0072】

また、クライアント端末1aの端末IDが「1」、クライアント端末1aの最大端末転送レートが「9600bps」、クライアント端末1aからリクエストされるURLが「http://www.a.co.jp/a.mpg」であることを前提とし、リクエストされたコンテンツの10秒分のデータをキャッシュ11に確保することを保証するものとする。

【0073】

リクエスト受付部4は、クライアント端末1aとのセッションを確立し、クライアント端末1aの端末ID「1」、最大端末転送レート「9600」、HTTPリクエスト「GET http://www.a.co.jp/a.mpg HTTP/1.1」を受信する。

【0074】

リクエスト受付部4は、キャッシュ11に「http://www.a.co.jp/a.mpg」が保存されていないため、リクエストURL「http://www.a.co.jp/a.mpg」をリクエスト転送部5に転送する。また、リクエスト受付部4は、クライアント端末1aの端末ID「1」、最大端末転送レート「9600」、リクエストURL「http://www.a.co.jp/a.mpg」をデータ送受信制御部7に通知する。

【0075】

データ送受信制御部7は、キャッシュ管理テーブル300に新たに項目を追加し、端末ID301に「1」を、URL302に「http://www.a.co.jp/a.mpg」を、最大端末転送レート304に「9600」を保持し、未送信バッファサイズの初期値を算出し、未送信バッファサイズ307に保持する(図10)。この未送信バッファサイズ307の算出方法は、最大端末転送レート「9600」で10秒間コンテンツを送出した場合に必要データサイズとして、「 $(9600/8) \times 10 = 12000 \text{ byte}$ 」となる。

【0076】

リクエスト受付部4からリクエストURL「http://www.a.co.jp/a.mpg」を転

送されたリクエスト転送部5は、「http://www.a.co.jp/a.mpg」のデータサイズを取得するために、リクエスト「HEAD http://www.a.co.jp/a.mpg HTTP/1.1」をWebサーバ13aに送信する。

【0077】

リクエスト転送部5は、Webサーバ13aとのセッションID「10」、Webサーバ13aへのリクエスト送信時刻「15000」をレスポンス監視部10に通知する。

【0078】

レスポンス監視部10は、リクエスト状況監視テーブル400に新たに項目を追加し、WebサーバセッションID401に「10」、リクエスト送信時刻402に「15000」を保持する(図6)。

【0079】

データ受信部6は、リクエスト転送部5がWebサーバ13aに送信したリクエスト「HEAD http://www.a.co.jp/a.mpg HTTP/1.1」に対するWebサーバ13aのレスポンスとして、「HTTP/1.1 200 OK Content-Length: 100000」を受信する。

【0080】

データ受信部6は、Webサーバ13aとのセッションID「10」、応答受信時刻「20000」、コンテンツサイズ「100000」をレスポンス監視部10に通知する。

【0081】

レスポンス監視部10は、レスポンス監視テーブル400のWebサーバセッションIDの一致する項目の応答受信時刻403に「20000」、コンテンツサイズ404に「100000」を保持する(図7)。また、レスポンス監視部10は、リクエスト送信時刻402「15000」と、応答受信時刻403「20000」の差分より、Webサーバ13aのレスポンス時間を「5000」と判定し、セッションID「10」、コンテンツサイズ「100000」、レスポンス時間「5000」をデータ送受信制御部7に通知する。

【0082】

データ送受信制御部 7 は、キャッシュ管理テーブル 300 の Web サーバセッション ID「10」と一致する項目の Web サーバレスポンス 306 に「5000」、コンテンツサイズ 310 に「100000」を保持し、未送信バッファサイズ 307 を更新する (図 11)。

【0083】

この時の未送信バッファサイズ 307 の算出方法は、Web サーバ 13a とのネットワーク遅延時間を含めるため、最大端末転送レート「9600」で、コンテンツ送出手間 10 秒と、Web サーバ 13a のレスポンス時間 5 秒の合計 15 秒間コンテンツを送出し続ける場合に必要データサイズとなるため、「 $(9600/8) \times (10+5) = 18000$ 」となる。

【0084】

データ送受信制御部 7 は、「http://www.a.co.jp/a.mpg」を 18000 byte 分取得するためのリクエスト「GET http://www.a.co.jp/a.mpg HTTP/1.1 Range: byte=0-17999」を Web サーバセッション ID「10」の Web サーバ 13a に送信するようにリクエスト転送部 5 に対して要求する。なお、このコンテンツ取得のリクエストは、適当なコンテンツサイズに分割して数回リクエストしてもよい。

【0085】

リクエスト転送部 5 は、データ送受信制御部 7 からのリクエストを Web サーバセッション ID「10」の Web サーバ 13a に送信する。リクエスト転送部 5 は、Web サーバ 13a とのセッション ID「10」、リクエスト送信時刻「20050」をレスポンス監視部 10 に通知する。

【0086】

レスポンス監視部 10 は、リクエスト状況監視テーブル 400 の Web サーバセッション ID 601「10」と一致する項目のリクエスト送信時間 402 に「20050」を保持し、応答受信時刻 403 をクリアする (図 8)。

【0087】

データ受信部 6 は、リクエスト転送部 5 が Web サーバ 13a に送信したリクエスト「GET http://www.a.co.jp/a.mpg HTTP/1.1 Range: byte=0-17999」に対

するWebサーバ13aのレスポンスとして、「HTTP/1.1 200 OK Content-Length: 18000 …」を受信する。

【0088】

データ受信部6は、受信したコンテンツをキャッシュ11に保持し、Webサーバ13aとのセッションID「10」、応答受信時刻「24000」、受信したコンテンツサイズ「18000」をレスポンス監視部10に通知する。

【0089】

レスポンス監視部10は、リクエスト状況監視テーブル400のWebサーバセッションID601「10」と一致する項目の応答受信時刻403に「24000」を保持する(図9)。データ受信部6は、リクエスト送信時間402「20050」と応答受信時刻403「24000」の差分の時間より、Webサーバ13aのレスポンス時間を「3950」と判定し、セッションID「10」、Webサーバのレスポンス時間「3950」、受信したコンテンツのデータサイズ「18000」をデータ送受信制御部7に通知する。

【0090】

データ送受信制御部7は、キャッシュ管理テーブル300のセッションID「10」と一致するの項目の未送信データサイズ308に「18000」、Webサーバレスポンス306に「3950」を保持し、未送信バッファサイズ307を更新する(図12)。

【0091】

この時の未送信バッファサイズ307の算出方法は、Webサーバ13aとのネットワーク遅延時間を含めるため、最大端末転送レート「9600」で、コンテンツ送出の保証時間10秒と、Webサーバ13aのレスポンス時間3.95秒の合計13.95秒間コンテンツを送出し続ける場合に必要なデータサイズとなるため、「 $(9600/8) \times (10 + 3.95) = 16740$ 」となる。

【0092】

この時点で、「未送信バッファサイズ307<未送信データサイズ308」の関係が満たされているため、端末ID「1」のクライアント端末1aに対し、Webサーバ13aのレスポンス時間を加味した13.95秒分のコンテンツを確

保できていることになる。このため、データ送受信制御部 7 はデータ送出部 8 に対し、キャッシュ 11 に途中まで保持しているコンテンツ「http://www.a.co.jp/a.mpg」を端末 ID「1」のクライアント端末 1a へ送出開始するよう要求する。

【0093】

データ送出部 8 は、端末 ID「1」のクライアント端末 1a に対し、キャッシュ 11 の「http://www.a.co.jp/a.mpg」を最大端末転送レート 9600bps で 1 秒分送出しようと試み、1 秒間で 1000byte のデータを送出できたとする。データ送出部 8 は、転送状況監視部 9 に端末 ID「1」、送出したデータサイズ「1000」、送出時間「1000」を通知し、引き続きデータ送出を行う。

【0094】

転送状況監視部 9 は、送出したデータサイズ「1000」、送出時間「1000」より、端末転送レートを 8000bps と判定し、データ送受信制御部 7 に端末 ID「1」、端末転送レート「8000」、送信データサイズ「1000」を通知する。

【0095】

データ送受信制御部 7 は、キャッシュ管理テーブル 300 の端末 ID「1」と一致する項目の端末転送レート 305 に「8000」、送信済みデータサイズ 309 に「1000」を保持し、未送信バッファサイズ 307 と、未送信データサイズ 308 を更新する。

【0096】

未送信バッファサイズ 307 は、10 秒分と、Web サーバレスポンス 306「3950」の合計 13.95 秒間、端末転送レート 305「8000」で送出するために必要なデータサイズ「 $(8000/8) \times (10 + 3.95) = 13950$ 」となり、未送信データサイズ 308 は、「 $18000 - 1000 = 17000$ 」を保持する（図 13）。

【0097】

データ送受信制御部 7 は、「未送信バッファサイズ 307 < 未送信データサイ

ズ 3 0 8」となっているため、WebサーバセッションID「10」に対する受信制御は行わないが、他のWebサーバとのセッションでのコンテンツ受信処理も行っていない場合、キャッシュ管理テーブルに登録している各Webサーバセッション毎にコンテンツ取得のリクエストを該当Webサーバに送信するようにリクエスト転送部5に要求する。

【0098】

以後、データ送出部8が、端末ID「1」のクライアント端末1aに対するデータ送出を繰り返し、データ送受信制御部7が、未送信バッファサイズ307と、未送信データサイズ308を更新した結果、「未送信バッファサイズ307<未送信データサイズ308」を満たさなくなった場合、データ送受信制御部7は、リクエスト転送部5に対し、「http://www.a.co.jp/a.mpg」の18000byte以降の続きのデータ取得のリクエストをWebサーバセッションID「10」のWebサーバ13aで送信するように要求する。

【0099】

上記のような、データ転送制御を行うことにより、端末転送レートの異なる複数のクライアント端末へのコンテンツ送出を効率的に行うことが可能となる。また、ネットワーク遅延時間の異なる複数のWebサーバからのコンテンツ取得を必要十分な負荷で行うことが可能となる。

【0100】

【発明の効果】

以上のように、本発明の一実施形態に係るデータ中継システムによれば、以下の効果を得ることができる。

【0101】

(1) クライアント端末に送信するコンテンツの残りデータサイズに余裕がないセッションを優先して、Webサーバからのコンテンツ受信処理を行うため、クライアント端末に対して、コンテンツを送出する際にデータ途切れが起こり難くなる。

【0102】

(2) クライアント端末に対し、一定時間送出し続けることができるだけのコ

ンテンツを確保してから、クライアント端末への送出を開始するため、コンテンツを送出する際にデータ途切れが起こり難くなる。

【0103】

(3) クライアント端末の転送レートにより、Webサーバからのコンテンツ受信処理の優先順位を制御するため、転送レートの速いクライアント端末にコンテンツを送出する際にも、データ途切れが起こり難くなる。

【0104】

(4) Webサーバからのコンテンツ受信処理の優先順位を制御する際に、クライアント端末のリアルタイムの転送レートをフィードバックするため、クライアント端末の転送レートの揺らぎが発生した場合にも、データ途切れが起こり難くなる。

【0105】

(5) Webサーバのレスポンス時間により、Webサーバからのコンテンツ受信処理の優先順位を制御するため、レスポンスの遅いWebサーバからのコンテンツをクライアント端末に送出する際に、データ途切れが起こり難くなる。

【0106】

(6) クライアント端末の転送レートで、一定時間送出し続けることが可能なコンテンツを確保してから、クライアント端末への送出を開始するため、クライアント端末の転送レートが揺らぐ場合でも、コンテンツを送出する際にデータ途切れが起こり難くなる。

【0107】

(7) クライアント端末の転送レートで、一定時間送出し続けることが可能なコンテンツに、Webサーバのレスポンス時間を含めたサイズのコンテンツを確保してから、クライアント端末への送出を開始するため、Webサーバのレスポンスが遅い場合でも、コンテンツを送出する際にデータ途切れが起こり難くなる。

【図面の簡単な説明】

【図1】

本発明の一実施形態に係るデータ中継システムの構成を示すブロック図

【図 2】

従来技術において、プロキシ・サーバでのデータ途切れが起こる場合の、Webサーバからの受信データサイズと、クライアント端末への送出データサイズの関係を示す図

【図 3】

本発明の一実施形態に係るデータ送受信制御部 7 が管理するキャッシュ管理テーブルを示す図

【図 4】

本発明の一実施形態に係るレスポンス監視部 10 のレスポンス監視テーブルを示す図

【図 5】

本発明の一実施形態に係るデータ中継システムのシーケンス図

【図 6】

本発明の一実施形態に係るレスポンス監視テーブルのある時点での内容を表す図

【図 7】

本発明の一実施形態に係るレスポンス監視テーブルのある時点での内容を表す図

【図 8】

本発明の一実施形態に係るレスポンス監視テーブルのある時点での内容を表す図

【図 9】

本発明の一実施形態に係るレスポンス監視テーブルのある時点での内容を表す図

【図 10】

本発明の一実施形態に係るキャッシュ管理テーブルのある時点での内容を表す図

【図 11】

本発明の一実施形態に係るキャッシュ管理テーブルのある時点での内容を表す

図

【図 1 2】

本発明の一実施形態に係るキャッシュ管理テーブルのある時点での内容を表す

図

【図 1 3】

本発明の一実施形態に係るキャッシュ管理テーブルのある時点での内容を表す

図

【図 1 4】

従来技術において、一般的なプロキシ・サーバの構成を示すブロック図

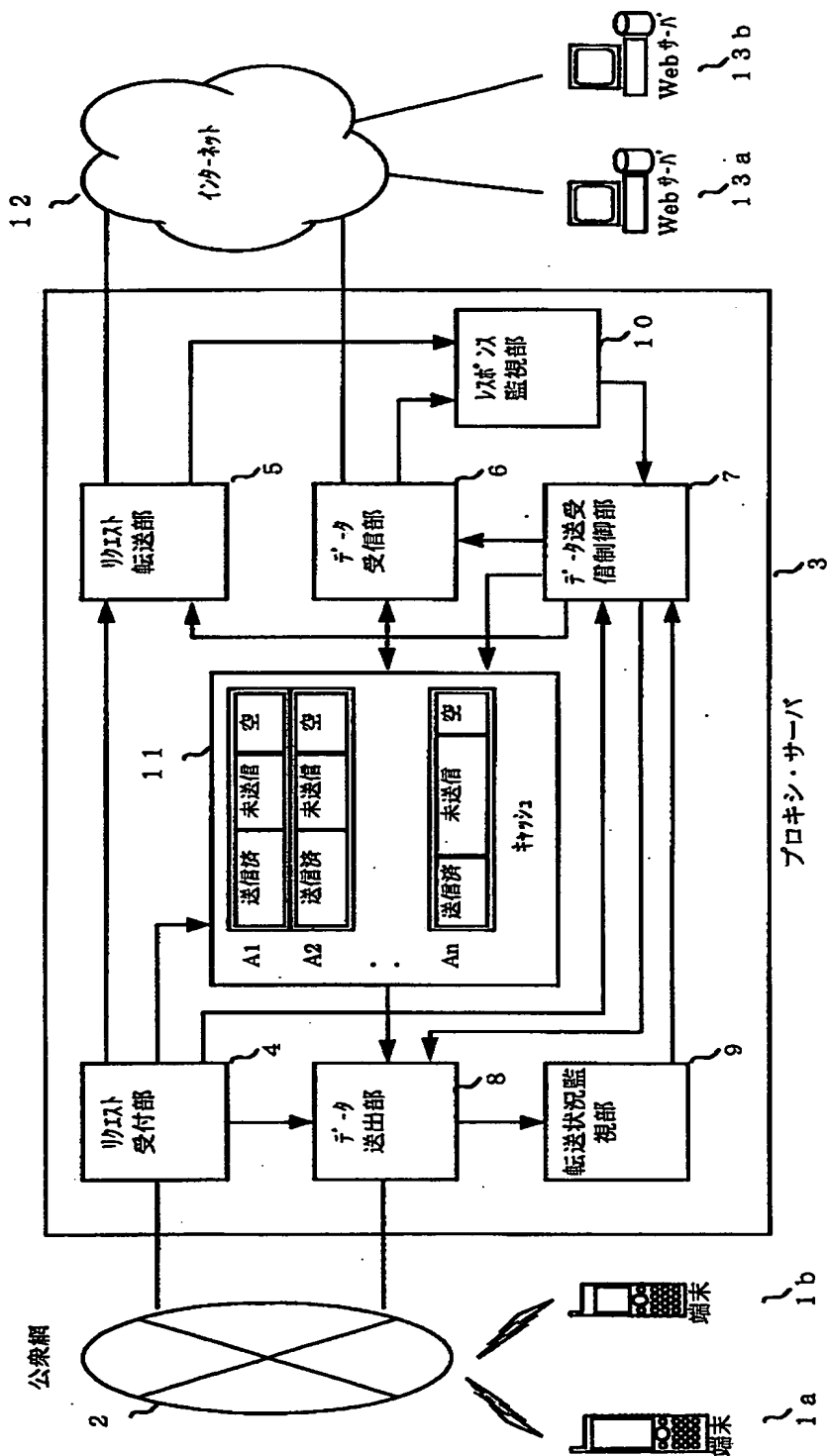
【符号の説明】

- 1 a クライアント端末 a
- 1 b クライアント端末 b
- 2 公衆網
- 3 プロキシ・サーバ
- 4 リクエスト受付部
- 5 リクエスト転送部
- 6 データ受信部
- 7 データ送受信制御部
- 8 データ送出部
- 9 転送状況監視部
- 10 レスポンス監視部
- 11 キャッシュ領域
- 12 インターネット
- 13 a Webサーバ a
- 13 b Webサーバ b
- 201 Webサーバから受信するストリーム・データサイズ
- 202 クライアント端末への送信済みストリーム・データサイズ
- 203 クライアント端末に未送信のストリーム・データ
- 300 キャッシュ管理テーブル

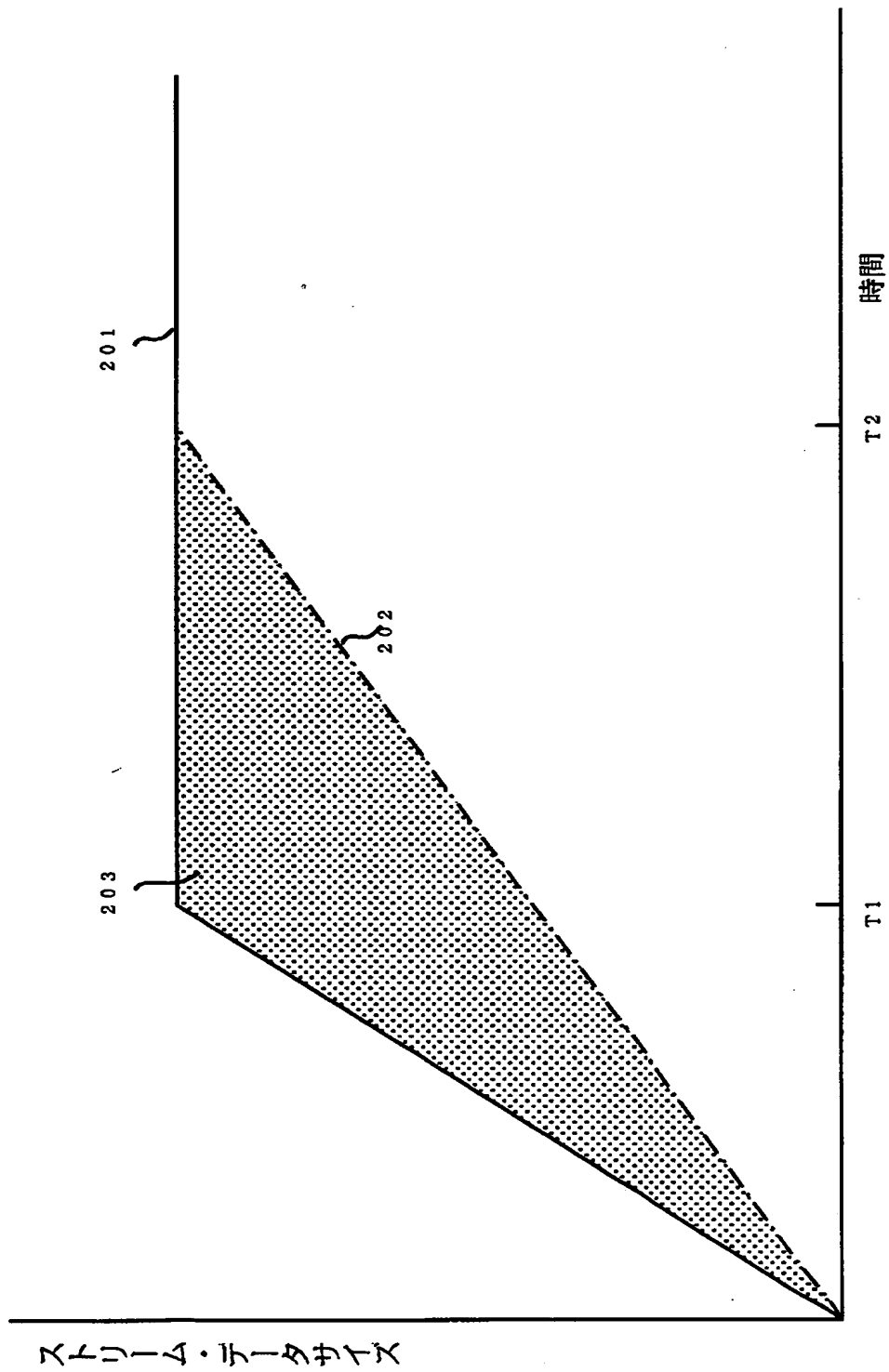
- 301 端末ID
- 302 URL
- 303 WebサーバセッションID
- 304 最大端末転送レート
- 305 端末転送レート
- 306 Webサーバレスポンス
- 307 未送信バッファサイズ
- 308 未送信データサイズ
- 309 送信済みデータサイズ
- 310 コンテンツサイズ
- 400 レスポンス監視テーブル
- 401 WebサーバセッションID
- 402 リクエスト送信時刻
- 403 応答受信時刻
- 404 コンテンツサイズ

【書類名】 図面

【図 1】



【図2】



【図3】

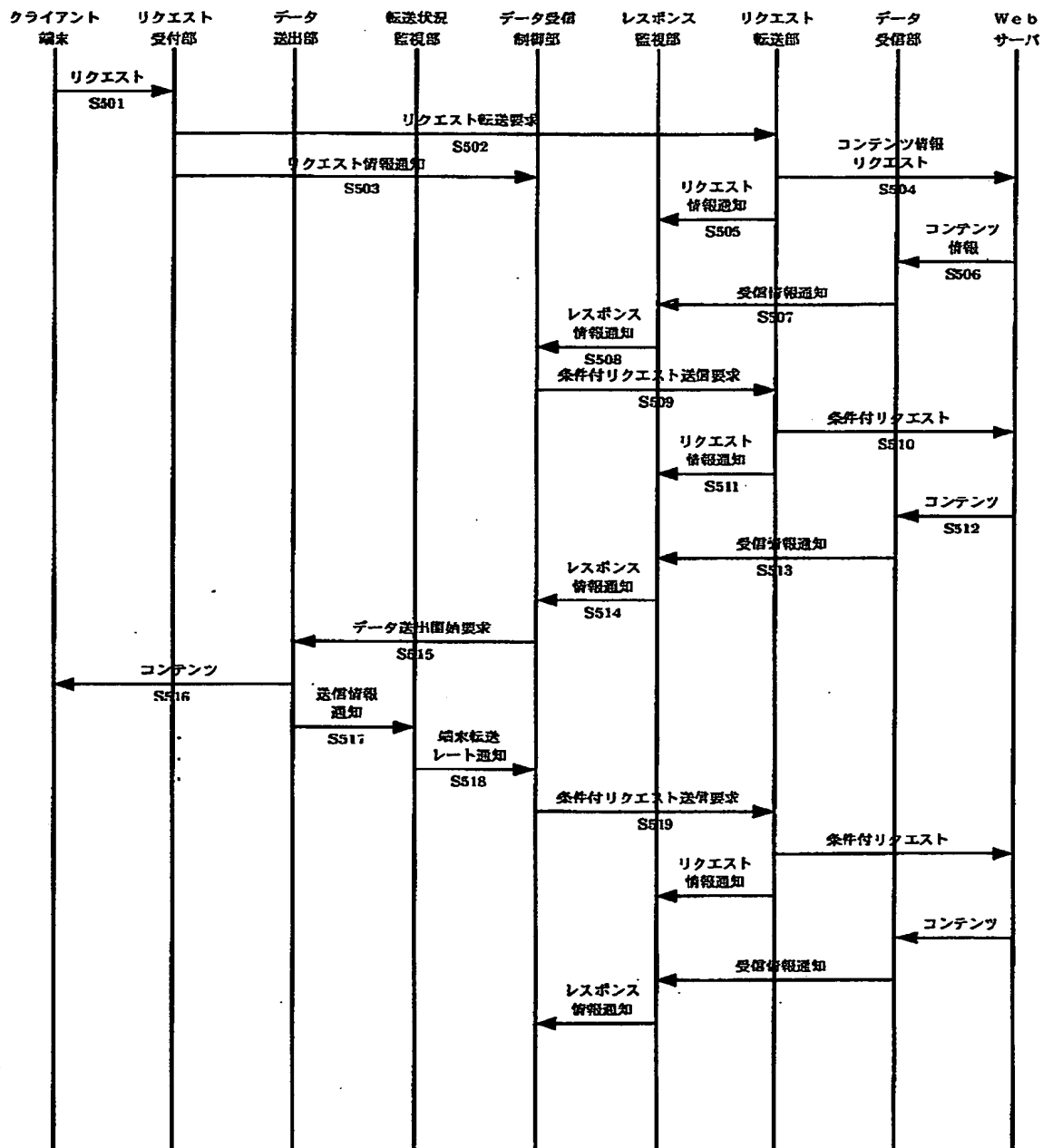
301	302	303	304	305	306	307	308	309	310
端末ID	URL	Webサーバ セッション ID	最大端末 転送レート (bps)	端末転送 レート (bps)	Webサーバ レスポンス (msec)	未送信 バッファ サイズ(byte)	未送信 データ サイズ(byte)	送信済み データ サイズ(byte)	コンテナー サイズ(byte)
1	http://www.a.co.jp/a.mpg	10	9600	3000	74000	3000	9750	10000	100000
2	http://www.a.co.jp/b.mpg	20	32000	20000	28000	29500	25000	20000	200000
...									

【図 4】

401	402	403	404
Web サーバ セッション ID	リクエスト 送信時刻 (msec)	応答 受信時刻 (msec)	コンテンツサイズ (byte)
10	15000	17000	100000
20	16000	18500	80000

400

【図 5】



【図 6】

Web サーバ セッション I D	リクエスト 送信時刻 (msec)	応答 受信時刻 (msec)	コンテンツ サイズ (byte)
10	15000	—	—
⋮			

【図 7】

Web サーバ セッション I D	リクエスト 送信時刻 (msec)	応答 受信時刻 (msec)	コンテンツ サイズ (byte)
10	15000	20000	100000
⋮			

【図 8】

Web サーバ セッション I D	リクエスト 送信時刻 (msec)	応答 受信時刻 (msec)	コンテンツ サイズ (byte)
10	20050	—	100000
⋮			

【図 9】

Web サーバ セッション I D	リクエスト 送信時刻 (msec)	応答 受信時刻 (msec)	コンテンツ サイズ (byte)
10	20050	24000	100000
⋮			

【図 1 0】

端末 ID	URL	Web サーバ セッション ID	最大端末 転送レート (bps)	端末転送 レート (bps)	Web サーバ レスポンス (msec)	未送信 パッファ サイズ(byte)	未送信 データ サイズ(byte)	送信済み データ サイズ(byte)	コンテンツ サイズ(byte)
1	http://www.a.co.jp/a.mpg	—	9600	—	—	12000	—	—	—
...									

【図 1 1】

端末 ID	URL	Web サーバ セッション I D	最大端末 転送レート (bps)	端末転送 レート (bps)	Web サーバ レスポンス (msec)	未送信 パッファ サイズ(byte)	未送信 データ サイズ(byte)	送信済み データ サイズ(byte)	コンデンス サイズ(byte)
1	http://www.a.co.jp/a.mpg	10	9600	—	5000	18000	—	—	100000

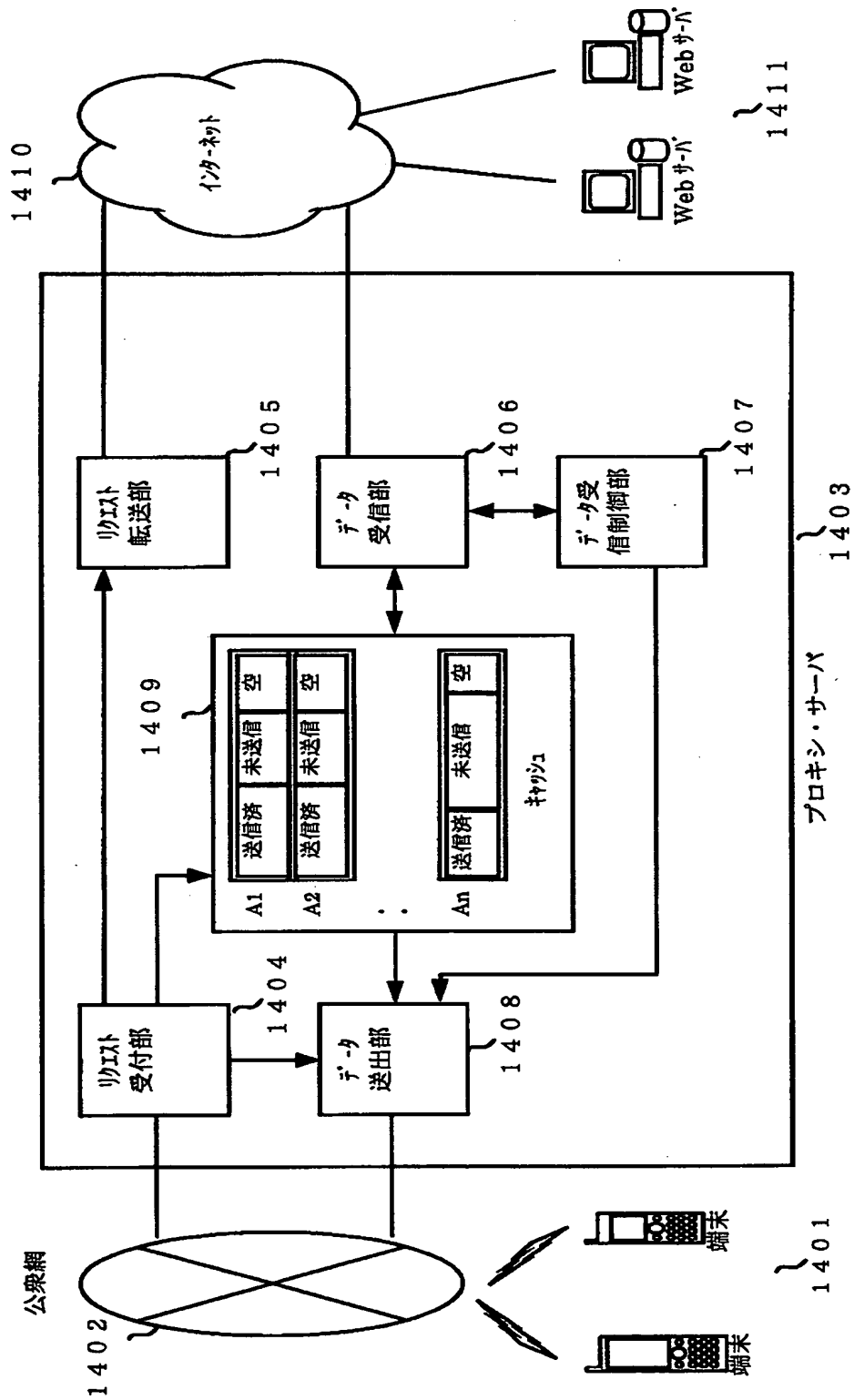
【図12】

端末ID	URL	WebセッションID	最大端末転送レート(bps)	端末転送レート(bps)	Webサーバーレスポンス(msec)	未送信パッファサイズ(byte)	未送信データサイズ(byte)	送信済みデータサイズ(byte)	コンテンツサイズ(byte)
1	http://www.a.co.jp/a.mpg	10	9800	—	3950	16740	18000	—	100000

【図 1 3】

端末 ID	URL	Web サーバ セッション ID	最大端末 転送レート (bps)	端末転送 レート (bps)	Web サーバ レスポンス (msec)	未送信 バッファ サイズ(byte)	未送信 データ サイズ(byte)	送信済み データ サイズ(byte)	コンテンツ サイズ(byte)
1	http://www.a.co.jp/a.mpg	10	9600	8000	3950	16740	17000	1000	100000
. . .									

【图 14】



【書類名】 要約書

【要約】

【課題】 インターネット上のストリームデータを転送する際のスループットの揺らぎに対して、データ途切れの起こりにくいシステムを提供することである。

【解決手段】 プロキシサーバ3は、クライアント端末1 aからのコンテンツ取得要求に該当するコンテンツがキャッシュ1 1に保存されていない場合、インターネット1 2を経由して、オリジナルコンテンツを保持しているWebサーバ1 3 aから取得し、クライアント端末1 aに送信する。この時、プロキシサーバ3は、Webサーバ1 3 aからのコンテンツ受信処理をレスポンス監視部1 0で監視し、クライアント端末1 aへのコンテンツ送信処理をデータ送出部8で監視する。ここで取得したWebサーバ1 3 aのレスポンス、および、クライアント端末1 aのスループットにより、Webサーバ1 3 aからのコンテンツ受信処理の優先順位をリアルタイムに変更し、データ途切れを起こりにくくする。

【選択図】 図1

出 願 人 履 歴 情 報

識別番号 [000005821]

1. 変更年月日	1990年 8月28日
[変更理由]	新規登録
住 所	大阪府門真市大字門真1006番地
氏 名	松下電器産業株式会社